Stimuli-Responsive Assembly of Hierarchical DNA Nanomaterials

Anjelica Kucinic

Undergraduate Thesis Defense

Nanoengineering and Biodesign Laboratory

The Ohio State University

Department of Chemical and Biomolecular Engineering

Faculty Advisor:

Dr. Carlos Castro, Department of Mechanical and Aerospace Engineering

Graduate Student Mentors:

Jenny Le, Biophysics Graduate Program

Chao-Min Huang, Department of Mechanical and Aerospace Engineering

Thesis Committee:

Dr. Carlos Castro, Department of Mechanical and Aerospace Engineering

Dr. Hai-Jun Su, Department of Mechanical and Aerospace Engineering

Acknowledgements

I would like to first and foremost thank Dr. Carlos Castro for the incredible opportunity to start with NBL as a member of OhioMOD 2015. Without the opportunity to conduct research in OhioMOD 2015, I would not be the researcher, mentor, and student I am today. I would also like to thank Dr. Hai-Jun Su for the opportunity to work with the DISL. I am forever grateful to Jenny Le for all her guidance and support throughout the past two and a half years, she has been an incredible mentor, coworker, and of course friend. I would like to express my sincerest gratitude to Chao-Min Huang for his mentorship this past summer as well as for his generosity, support, and teaching me a little bit about mechanical engineering fundamentals.

I would also like to express my gratitude to Patrick Halley (OhioMOD 2016 graduate mentor) for his creativity and youthful spirit which helped OhioMOD 2016 achieve such great success. I would like to express my gratitude to Joshua Johnson for his mentorship in OhioMOD 2015 as well as Dr. Chris Lucas and other NBL graduate mentors for all their support and company in the lab. Many thanks to Melika Shahhosseini for making sure I eat homecooked meals when needed. Members from the Chemical Engineering department, specifically Dr. John Clay and my study group have helped keep me sane throughout this process. My family and friends have been incredibly understanding of my time commitment to this meaningful research, which I appreciate beyond words.

Last but never least, I would like to thank the NBL undergraduates past and present, especially the members of OhioMOD 2016 for the endless laughter, memories, and friendship. This group is very special and I will cherish the memories we made together, from trips to Cazuela's and Buckeye Donuts to late-night research sessions that end with nonsensical life talks.

Abstract

DNA origami is an emerging nanotechnology for fabrication of nanostructures with promising applications in biosensing, drug delivery, and nanomanufacturing. This approach uses hundreds of oligonucleotides (short pieces of single-stranded DNA, ssDNA) to fold a long ssDNA "scaffold" strand into precisely designed nanoscale 3D geometries. Our laboratory is interested in making responsive materials and devices for applications including drug delivery, biosensing, and biophysical measurements. This work will develop a framework for creating stimuli-responsive DNA material assemblies based on dynamic structures with "cryptic" or hidden binding sites that are occluded and only become accessible for material assembly after reconfiguration of the structure. We will utilize complex multiple arm structures connected by joints that are initially closed to occlude material assembly sites, and opens in response to specific trigger molecules. The actuation or triggering of these devices may result in a hierarchical structural assembly, allowing for a larger surface area with more exposure to harmful proteins or molecules. The triggered hierarchical assembly of DNA nanostructures has promising applications such as diagnostic tools, delivery systems, or capture devices.

Contents

List of Figures

Chapter 1: Introduction

1.1: Motivation for Diagnostic Tools

Diagnosis of many diseases, such as Alzheimer's disease (AD) (1), is typically late. With AD, diagnosis often occurs after symptoms such as the inability to remember newly developed memories or functions, and the aggregation of beta-amyloids is already blocking communication among nerve cells thus disrupting normal cellular processes (1). Current methods of diagnosis rely heavily on monitoring brain function and mental decline with the use of computerized axial tomography (CAT) scans and lumbar punctures (spinal taps) (1). Although these detection methods are minimally invasive, they are not capable of detecting beta-amyloid plaques prior to aggregation. Detecting beta-amyloids before the onset of aggregation, and before the destruction of brain size and function, can potentially prevent further progression of the disease and extend the quality of life for each diagnosis. Early detection may lead to prevention of the progression of AD as well as effective treatment plans. This work focuses on the development of a new sensing approach based on DNA nanotechnology, specifically a DNA origami hierarchical assembly system.

1.2: Introduction to DNA Origami

DNA origami is an emerging field in nanotechnology. The field was pioneered by Paul Rothemund and consists of the molecular self-assembly of complex 2D and 3D DNA nanostructures. The structures are formed from a long, circular single-stranded DNA (ssDNA) called the "scaffold" and multiple short DNA oligonucleotides called "staples" that are complementary to the scaffold strand in a piecewise manner (3). The scaffold used in this lab is derived from the M13mp18 viral genome and is typically between 7000 to 8000 base pairs (bp) long. When the scaffold and staples are mixed together and heated, the staples form double-

stranded (ds) regions on different sections of the scaffold, thus “pinching” together parts of the scaffold to create a desired geometry. In the self-assembly fabrication process (10), DNA interactions are initially melted at a high temperature. Once melted, the structures are set to an annealing temperature to allow for binding of staples to the correct locations on the scaffold. The folded DNA origami structures are then cooled to room temperature and then typically stored a 4^{o}C. The desired geometries are designed through a program called caDNAno (4). The scaffold routing in caDNAno has Holliday junctions which are cross-shaped structures (Figure 1) that also form during genetic recombination, producing four strands of DNA capable of exchanging genetic information (5,6,16).

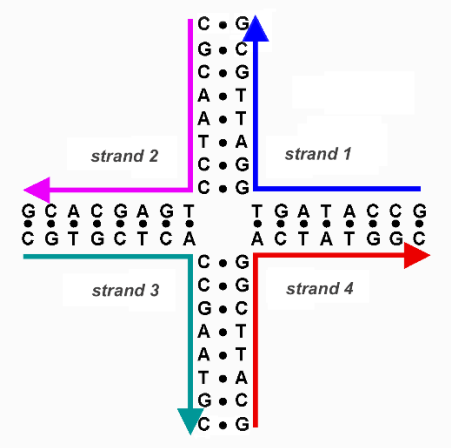


Figure 1: Holliday Junction, cross-shaped structure containing four ssDNA molecules (11).

The Holliday junctions help to bind individual ds helices of DNA together into a larger geometry. These geometries are programmed based on the staple sequences that are complementary to the scaffold itself. DNA origami geometries include 2D structures such as smiley faces and triangles as well as 3D structures modeling different joints such as a revolute joint, known as the hinge (3,7).

1.3: Background of Triggered Assembly

Triggered assembly of DNA nanostructures have the potential to create hierarchical structures that can function as diagnostic or delivery systems for treatment and prevention of different diseases. Since DNA origami structural dimensions are limited by the scaffold strand, a strategy where creating tile frameworks of DNA origami introduces larger structures (2). Figure 2 depicts a process in which a pre-formed scaffold frame is used along with DNA origami tiles to create a larger structure called Super Origami.

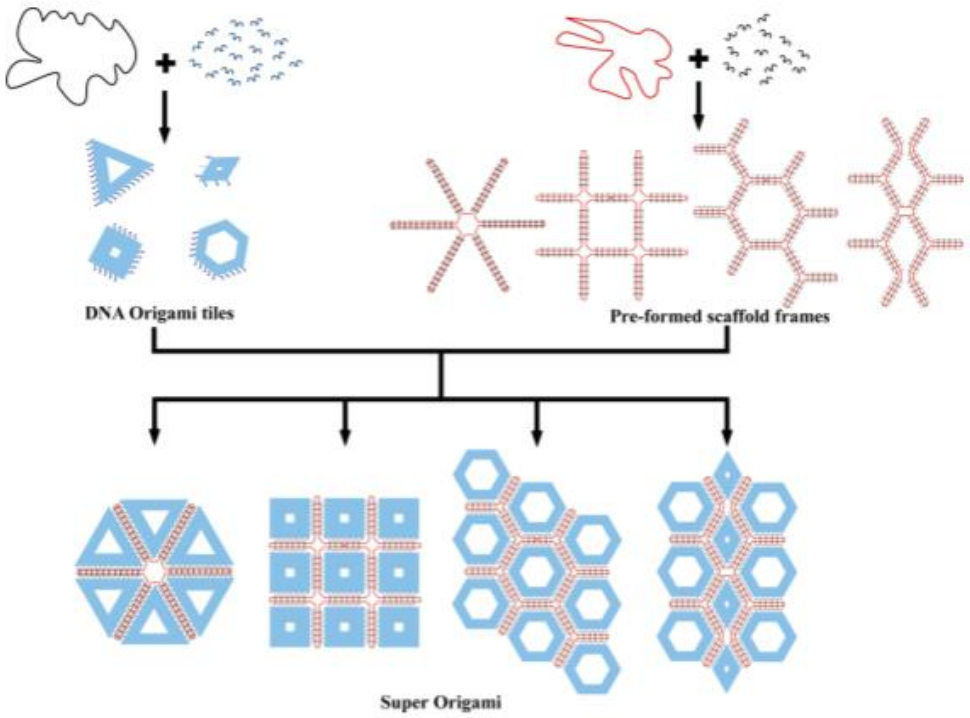


Figure 2: Organizing DNA Origami Tiles into Larger Structures Using Preformed Scaffold Frames (2)

Expanding and adjusting the size of DNA origami has been difficult, due to the limitations set by the scaffold. However, introducing trigger molecules to polymerize one DNA nanostructure, allows for the assembly of larger frameworks with potential for real-time control over material assembly or the possibility for responsive assembly that is triggered by changes in the local

environment. Triggering a nanostructure to change conformations via a specified protein such as beta-amyloids, with the use of protein- specific aptamer strands is a mechanism that potentially establishes a DNA nanostructure diagnostic tool.

One method used to actuate or trigger monomer (single) DNA nanostructures is called strand displacement. An overhang is a dsDNA strand of specific sequence placed on one or multiple site on a DNA nanostructure, and can have a complementary ssDNA strand for displacement. The complementary ssDNA sequence can be referred to as actuating strands and can trigger an opening or closing mechanism of a nanostructure (12). Furthermore, a toehold is an extension of unpaired bases, typically around five single-stranded bases, that overhang or from a dsDNA duplex. Displacement strands are strands of DNA that are complementary to a full strand including the toehold region. After first attaching to the toehold region, the actuation strand can displace the shorter strand in the DNA duplex. Strand displacement complexes or cascades can be used to assemble dynamic nanostructures. The advantage of strand displacement complexes and cascades to assemble dynamic nanostructures is the isothermal temperature setting. Thermodynamic approaches require a thermal annealing step, or at least a temperature change, to trigger an assembly, which can make biological applications of the dynamic assemblies more difficult (17).

1.4: Background and Significance of SLUG, 4-Bar, SLL

With the use of caDNAno and DNA origami, the design for selective latching universal gadget (SLUG), Straight-Line Linkage (SLL) and transformable 4-Bar Linkage (4-Bar) were inspired by earlier mechanisms developed in our lab (7,18).

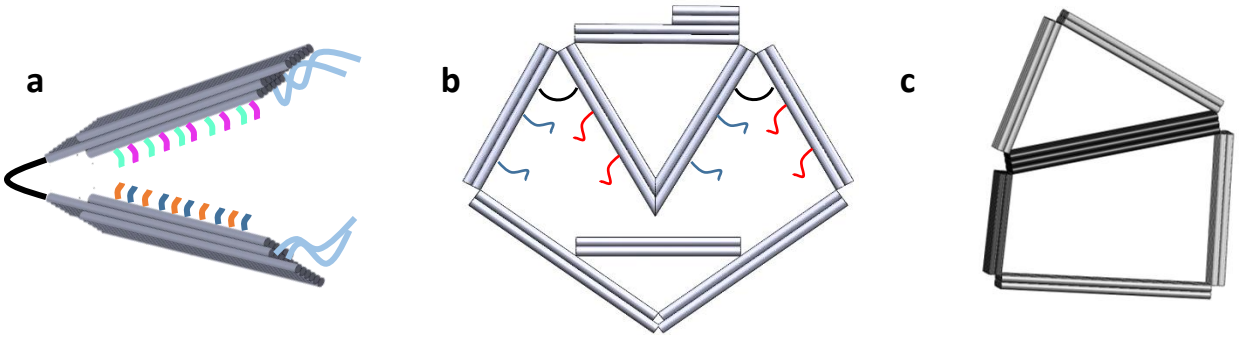


Figure 3: (a) SLUG design including internal overhangs, latching strands, and 6 30nt arm connections giving a 360^{o} motion. (b) SLL design including a symmetric structure with overhang sites on 4 bundles allowing for actuated motion of the triangle tip. (c) 4-Bar with bundles that are breakable to change conformations.

The SLUG is a simple monomer design that is similar to a previously developed hinge (7,19) however, it has six connections of equal length between the two arms. The SLUG was designed to have 3 different monomers with varying cryptic internal overhang sequences that are complementary to an arm on a different SLUG. This design allows for the controlled polymerization of the SLUG, once the cryptic overhangs are revealed. The Straight-Line Linkage consists of eight bundles or "arms" that are connected by ssDNA connections that are both long (30nt) and short (2nt) hinge connections. The axis of rotation is defined by the short connections. The structure has a symmetric design with two bundles hanging in the top middle to create a triangle tip. The A configuration of the 4-bar design consists of a triangular top (3 – 6 helix bundles) and rectangular ground or bottom (4 – 6 helix bundles) that has ssDNA connections at each vertex and toe-holds on each bundle, allowing for displacement of bundles. The displacement of these bundles allows for the transformation to three other variations of the 4-Bar, contributing to a transformation mechanism.

The SLUG design shows that triggering a single structure to polymerize can be used as a diagnostic tool, and the actuation of the SLL can be used as a detection device. The conformational changes characteristic with the 4-Bar can be used to fit differently shaped proteins such as a capture device.

1.5: Thesis Objectives and Hypothesis

The work in this thesis aims to provide a proof of concept for a diagnostic and delivery tool with the SLUG, as well as proof of concept and validation with the design and characterization of the SLL and 4-Bar Linkage. This work also proves to develop a framework of different actuating methods and polymerization via actuation.

1.6: Thesis Overview

The prevalence of Alzheimer ’s disease and other neurodegenerative disease continue to be an issue in the biomedical field, where diagnostic and therapeutic tools are limited to inefficient methods that do not prevent or cure the progression of these terminal diseases. With the use of DNA nanotechnology, specifically DNA origami, there is a promising potential for diagnostic, therapeutic and delivery devices. The design of the SLUG, SLL, and 4-Bar are motivated with the idea of triggered assembly and actuation mechanisms. The use of triggered assemblies can further create hierarchical structures for future use in the nanotechnology field as diagnostic tools, delivery mechanisms, and therapeutic devices.

Chapter 2: Selective Latching Universal Gadget (SLUG)

2.1: SLUG design

The Selective Latching Universal Gadget (SLUG) was designed by OhioMOD 2015, using caDNAno, based on the hinge created by Marras (7). The caDNAno design of the SLUG is depicted in Figure 4 below.

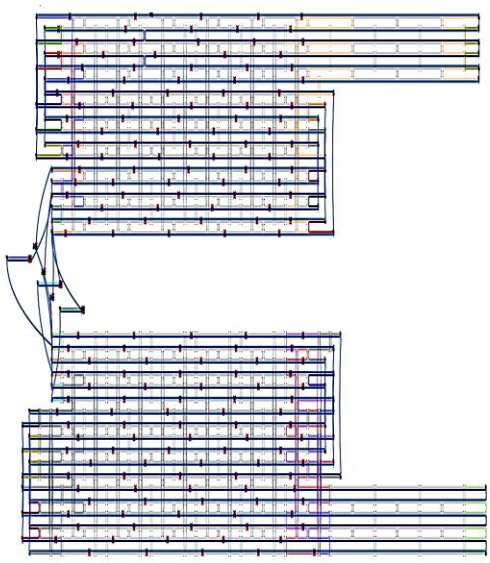


Figure 4: caDNAno design of the SLUG with two 3x6 square-lattice arms connected by 6 (30nt) ssDNA connections.

The structure consists of two 3x6 arms (18 helix bundles in a square lattice configuration), with six ssDNA connections (30nt) between the arms. The inner layer on both arms has 10 overhang sites, dsDNA strands placed in the scaffold routing, used as closing/assembly overhangs on the inner faces. There are three different SLUG monomers that each have different closing/assembly overhang sequences on the inner faces. The closing/assembly overhangs on the inner faces are complementary to closing/assembly overhangs on the inner faces of different SLUG monomers. Figure 5 shows a Solidworks schematic of a SLUG monomer.

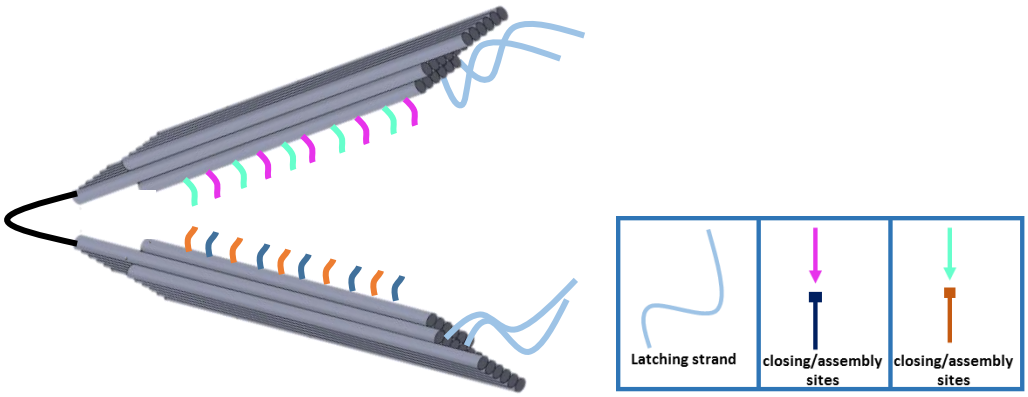


Figure 5: SolidWorks Schematic of SLUG with internal overhangs, arm connection, and latching strands.

There are three versions of the SLUG monomer, each with 2 different pairs of complementary overhang sequences, making for 4 different overhang sequences for each SLUG monomer. On the middle layer of each arm, on the opposite end of the hinge connections, there are two latching overhang sites (30nt) for latching.

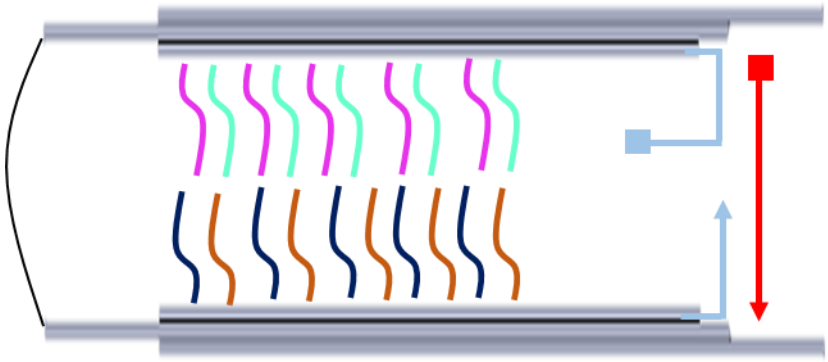


Figure 6: SolidWorks schematic of closed SLUG monomer ssDNA strand complementary to latching strand.

Figure 6 shows the closed SLUG monomer with a closing scheme involving the latching strand.

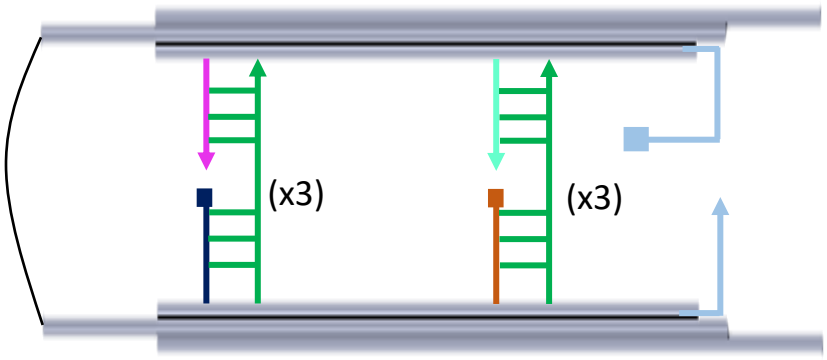


Figure 7: Closed SLUG monomer via complementary ssDNA low-affinity internal binding staples.

Figure 7 shows a closed SLUG monomer with a closing scheme involving ssDNA low-affinity closing/assembly binding staples.

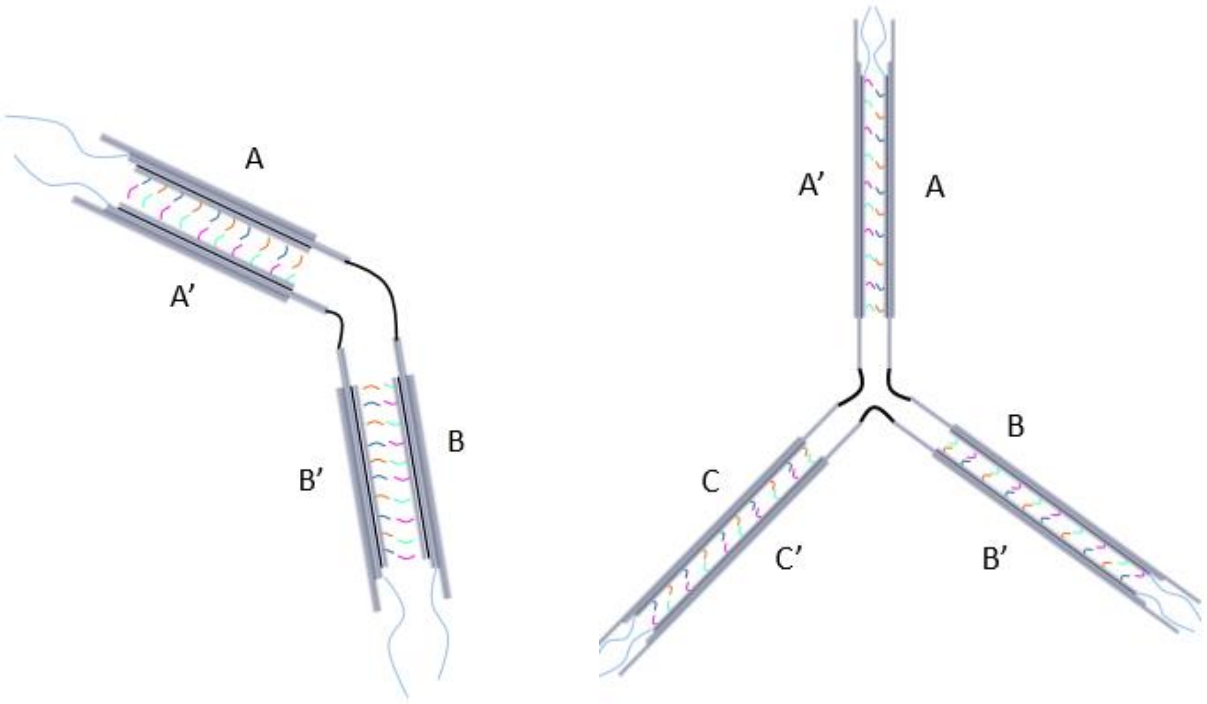


Figure 8: (left) Dimer and (right) Trimer of SLUG.

The polymerization of the SLUG focuses on the idea of creating a hierarchical structure that can be used as a net. The SLUG is designed in such a way that it can polymerize into a dimer structure or a trimer structure. Figure 8 shows SolidWorks schematics of the dimer and trimer SLUG.

2.2: Characterization and Validation of SLUG design

The SLUG structure is characterized by a 4-hour thermal annealing folding reaction (Chapter 5 section 1) and salt conditions. The 4-hour folding ramp consists of a 15-minute interval for melting the DNA at 65^{o}C, a 4-hour interval at the thermal annealing temperature, and a 15-minute cool-down at 4^{o}C to keep the DNA in a well-folded stable condition. The SLUG folds at an annealing temperature of 54.5^{o}C and 18 mM $MgCl_2$.

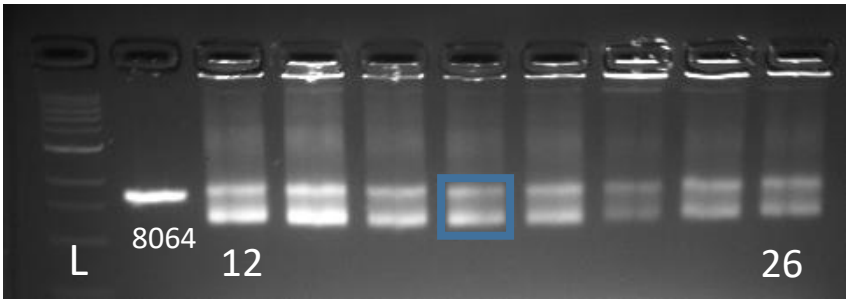


Figure 9: left to right: 1 kilobase ladder, scaffold, SLUG folded with 12mM-26mM $MgCl_2$ in increments of 2mM. The blue boxed lane indicated a well folded structure at a concentration of 18mM $MgCl_2$.

From Figure 6, well-folded structure shows up as a band that runs faster than the scaffold band. The magnesium folding screen in Figure 9 determines the optimal salt conditions for a DNA nanostructure. A 4-hour 56-52^{o}C thermal folding ramp with optimal salt conditions, shown in Figure 10, is run to determine the thermal annealing temperature. In Figures 9 and 10, the slug structure exhibited two bands, which could be indicative of multiple populations. However, there was no obvious difference between these populations when imaged by transmission electron microscopy.

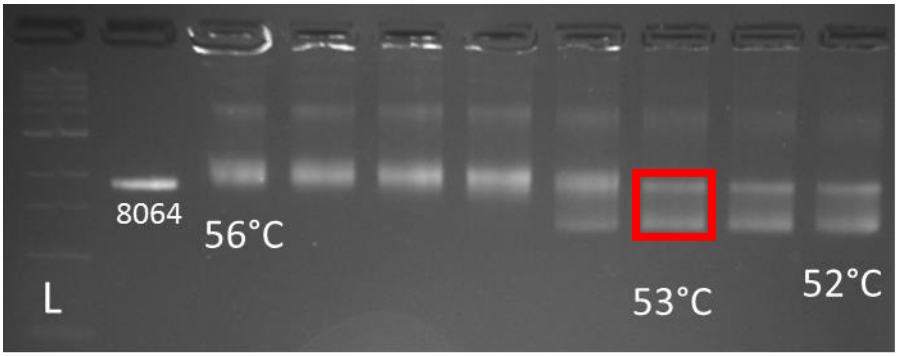


Figure 10: 4-hour thermal annealing fold to determine thermal annealing temperature for the SLUG at optimal salt conditions. The red box indicates that the SLUG annealing temperature is about 53°C. This is shown from the band shift present in the lane next to the 53°C lane.

After agarose gel electrophoresis verification of well-folded monomers, structures are imaged and validated with the use of Transmission Electron Microscopy (TEM). Figure 11 shows sample TEM images.

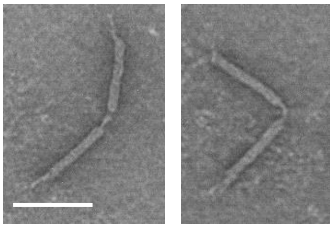


Figure 11: TEM images of well folded SLUG. Scale bar is 80 nm.

2.3: Opening and Closing SLUG Mechanisms

The SLUG is designed in such a way that the monomer can be closed shut via the latching strands, or via the internal overhangs. The idea for the closing and opening mechanism is to potentially hide a payload among the internal overhangs and then reveal that payload and internal overhang sequences, or to occlude strands that are responsible for higher order assembly. Here we focused on the higher order assembly, where revealing the internal overhangs allows for the three different monomers to polymerize, thus creating a hierarchical structure.

The closing mechanism consists of well-folded open SLUG monomers combined an excess of ssDNA strands that are complimentary to the latching overhangs. The intention is to “latch”

the monomers in a closed configuration and hide the internal sides of the monomer arms. However, due to the low binding affinity of the ssDNA closing staple, the efficiency for this mechanism is close to 10%. Figure 6 depicts the closing mechanism of the ssDNA closing strands and latching strands.

The internal overhang closing mechanism starts with a combination of opened, well-folded monomer, and 100 times excess ssDNA low-affinity internal binding staples that are complementary to the internal overhangs. This mechanism closes the SLUG, where a closed SLUG can be seen in figure 7.

Example images from the internal binding staple closing mechanism are depicted in figure 12 via TEM.

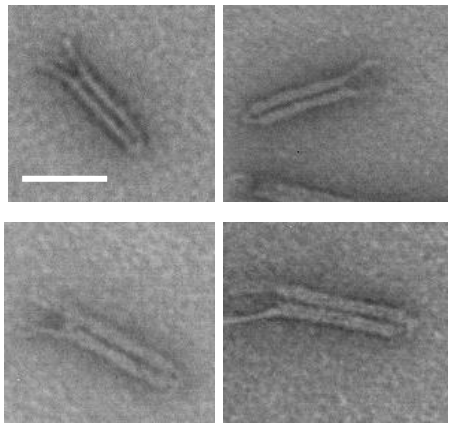


Figure 12: TEM images of closed SLUG monomers via complementary ssDNA low-affinity internal binding staples. Scale bar 100 nm.

Since there are three different SLUG monomers, the GC base-pairing content of the internal binding staples are different, thus showing varying degrees of closing efficiency. From figure 13, the closed monomer CA' has a closing efficiency of about 87% with a GC content of 40.5%. The monomer AB has a closing efficiency of about 89% with a GC content of 43.8%. The

highest closing efficiency is seen with closed monomer B'C' at 96% efficiency with a GC content of 53.1%. This data further indicates that a higher GC content has a higher binding affinity to close the monomer.

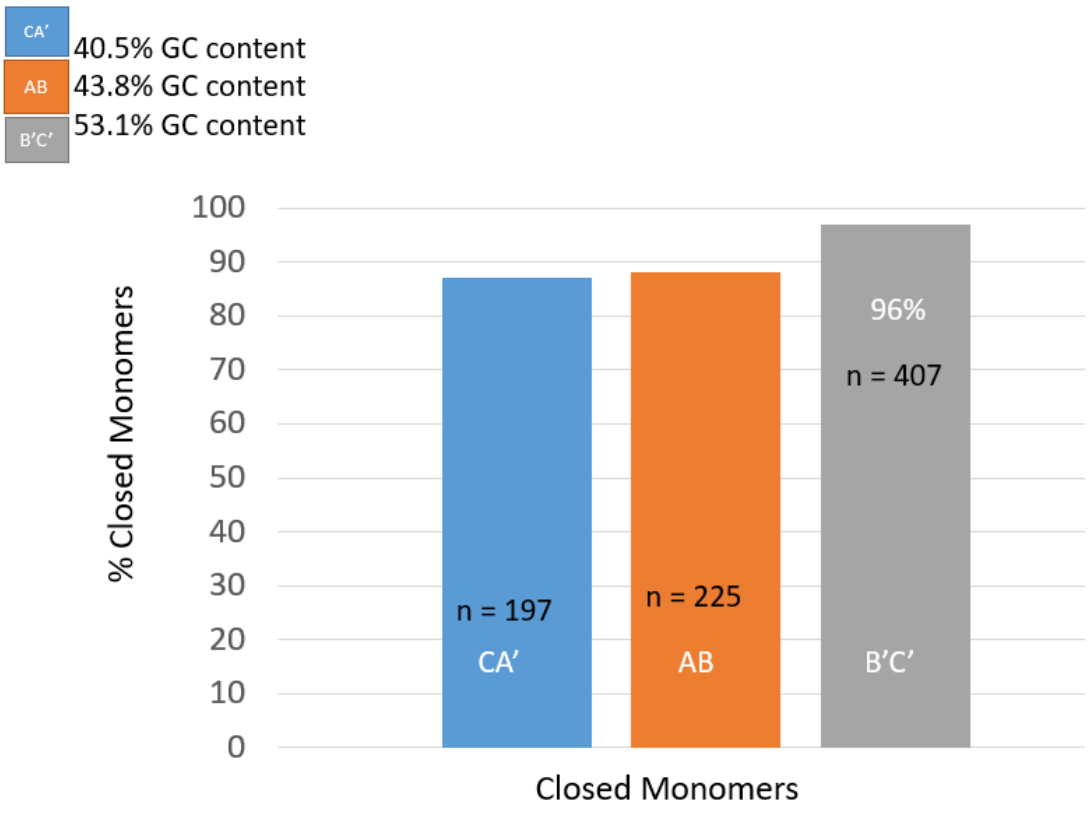


Figure 13: Closing Mechanism Efficiency with Internal Binding Staples

The opening mechanism for the SLUG is necessary to reveal the internal binding overhangs and potential therapeutic payload. The opening mechanism was validated with the use of TEM imaging, shown in figure 14. These sets of images were also used to calculate the opening mechanism efficiency for each SLUG monomer.

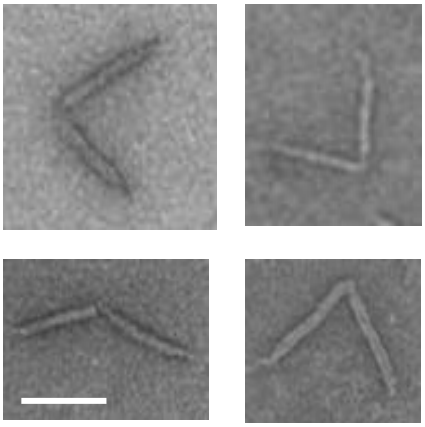


Figure 14: TEM images of opened SLUG monomers via opening mechanism. Scale bar 80nm.

The opening mechanism starts with closed SLUG monomers that were closed via internal binding staple mechanism, followed by polyethylene glycol purification process (20). The polyethylene glycol purification (PEG) method (Chapter 5 section 2) is used to "wash" away the internal binding staples, to actuate the opening configuration of the SLUG. Due to the variations in SLUG internal binding staple sequences and GC content, the efficiency of the opening mechanism varies depending on each version of the SLUG. The "washing" process was repeated three times, and opening efficiencies were measured for each version of the SLUG as shown in Figure 15.

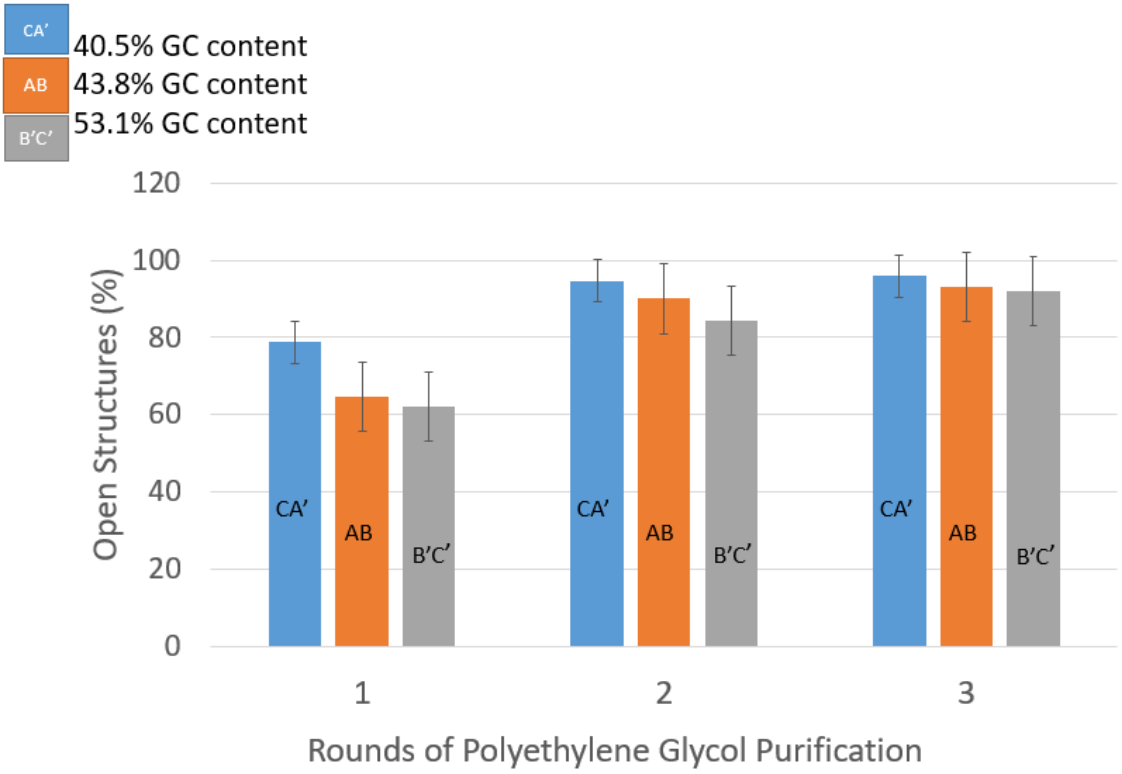


Figure 15: Opening Efficiency of SLUG monomers. The higher the GC content (%) the lower the efficiency of opening the monomer.

Figure 15 shows that after each round of "washing" the internal binding staples away, the higher the fraction of open monomers. After three rounds of PEG purification, open monomer CA' with a GC content of 40.5% has an opening efficiency of about 96%. Open monomers AB and B'C' with GC contents 43.8% and 53.1%, respectively, have opening efficiencies of about 93% each after three rounds of PEG purification. The figure also further suggests that the higher the GC content of the internal binding staples, the less likely the monomers are to open.

The opening and closing mechanisms of the SLUG are necessary for the polymerization of the SLUG, capable of capturing proteins or harmful molecules.

2.4: Higher order assembly of SLUG

Hierarchical assembly of the SLUG is effective when the closing and opening mechanism of monomer is effective and efficient. Starting as a smaller structure in solution, then triggering the structure to form higher order assembly has several potential advantages. For example, for biological applications, having a smaller structure enter the blood stream could help avoid triggering an immune response, and then forming a hierarchical structure that has a larger interaction area could potentially more easily capture or detect a protein.

Figure 16 shows the experimental results from the Trimer formation of directly folded SLUG monomers. These trimers were formed by incubating at 37^{o}C for 24 hours with equimolar concentrations of the three SLUG monomers. The trimers formed at an efficiency of 60%.

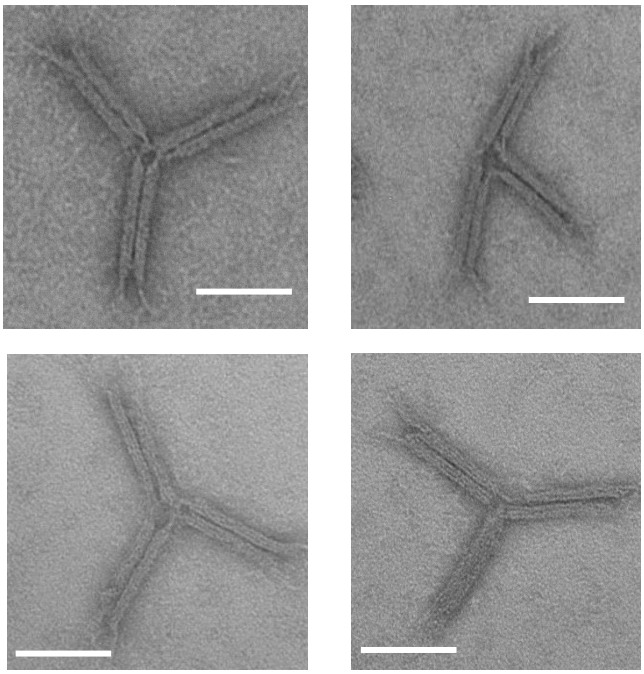


Figure 16: Trimer Formation of directly folded SLUG monomers. Scale bars = 80nm.

Chapter 3: Straight-Line Linkage (SLL)

3.1: SLL design and Motivation

The Straight-Line Linkage (SLL) is a structure designed by Mechanical engineering PhD student Chao-Min Huang. The structure (Figure 17) is designed so the coupled motion of several revolute joints (hinge) results in a motion path that approximates a straight line for one point on the structure (P3 in figure 17b). The SLL is composed of 8 bundles with 4x2 cross-sections with square lattices. The bottom truss has a 3x2 cross section. The top triangle and the ground link (l1) are connected to the input link (l2) and the output link (l4), shown in figure 17a, with two short connections and two long connections like the hinge design (7). The structure itself is symmetric in design, with a 22-base pair (bp) marker on one side of the in Figure 17a, b. The purpose for the marker is to distinguish between sides when the SLL is actuated, discussed in section 3.2. The SLL consists of 8 bundles that are connected by ssDNA pieces. Since the design of 3D multi-bundle structures in caDNAno is full of forced connections contributing to an increase in human error, the simulation package OxDNA (21) is employed to validate the routing. OxDNA also provides a way to compare simulations results with experimental results.

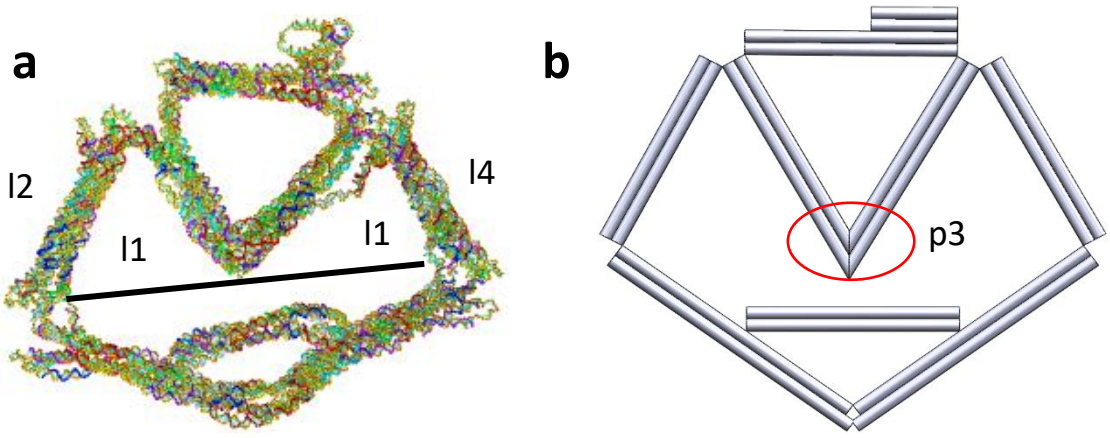


Figure 17: (a) OxDNA simulation of SLL with dsDNA helices visible in each bundle. (b) SolidWorks schematic of the SLL with 8 visible arms connected by revolute joints. The middle section containing the triangle portion, contains a triangle tip p3 that is used to measure actuation capabilities.

A well folded SLL was characterized with a folding reaction for 2.5 days and validation of well folded structure with agarose gel electrophoresis.

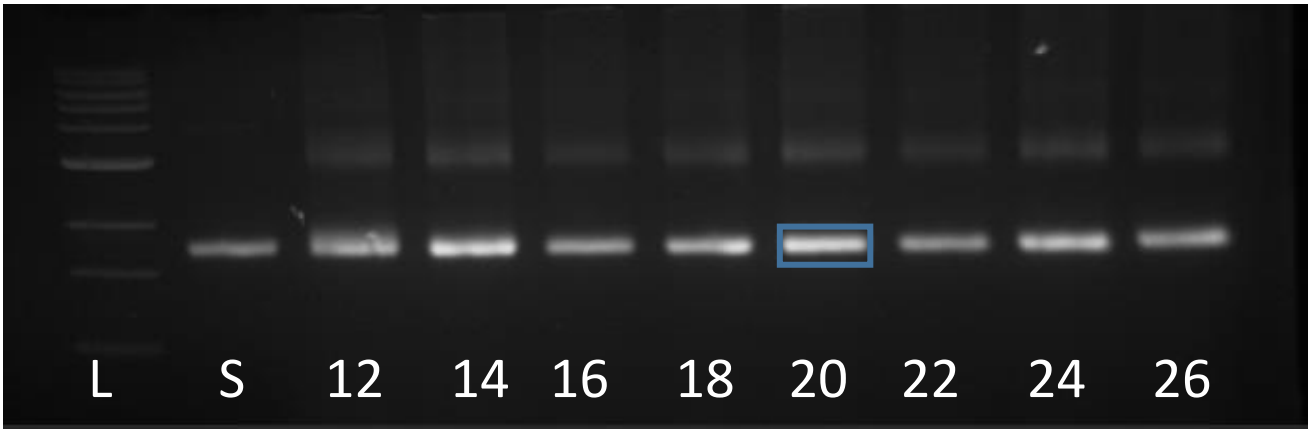


Figure 18: Gel of directly folded SLL with a $MgCl_2$ (mM) screen. From left to right starts with 1kb ladder, 7560 base scaffold (S), and $MgCl_2$ (mM) concentration gradient. SLL is characterized with a well-folded structure at 20mM $MgCl_2$ (blue box lane), where the lane is running at a speed similar to the scaffold (S) band.

The SLL models multiple hinge motions, making the structure more complex and difficult to measure the actuation capabilities and range of motion of the triangle tip. Figure 19 shows the lengths of each bundle in the top portion of the SLL and P3, which is the tip of the triangular portion.

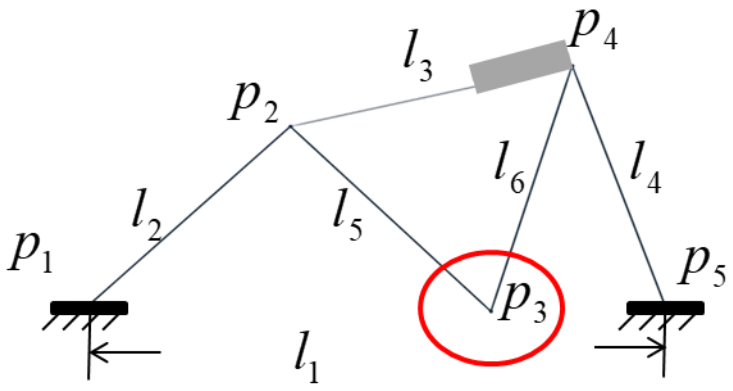


Figure 19: Top of SLL with Triangle Tip, p3. From p1 to p2 is length 2 (l2), from p2 to p4 is length 3 (l3), from p4 to p5 is length 4 (l4), from p2 to p3 and p4 to p3 is length 5 (l5) and length 6 (l6), respectively. From p1 to p5 there is a break from the bottom portion of the SLL, allowing for length 1 (l1). The length values can be measured from simulations and experimental data (TEM), which can be used to geometrically measure the angle of p3. P3 is the point at which the SLL is actuated.

The kinematic motion of p3 from Figure 19, is measured where 5 points (p1, p2, p3, p4, p5) are manually measured within a custom MATLAB code (Appendix C) written by Chao-Min Huang. The length of the bundles are directly calculated, and the position of the point p3 is also determined. Solving for the position of p3 results in measurement data that can be used to analyze the range of motion of the actuation point (p3) for the SLL. An example of the measurement points on a TEM image of a directly folded structure is shown in Figure 20.

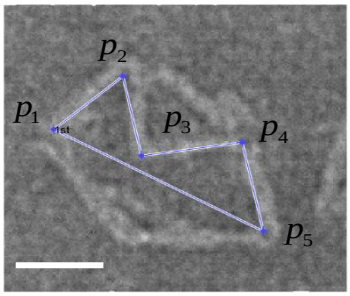


Figure 20: TEM image processed with the use of a Trigonometric MATLAB code that collects the lengths between p1,p2,p3,p4, and p5. Scale bar is 100nm.

In Figure 21 the kinematic motion of the unactuated SLL is shown. A trigonometric solver is used to get the kinematic curve (blue line). The lengths of each bundle are variables in the code that come directly from design values. Experimental data from TEM images is also plotted along the kinematic curve to show the experimental motion of the unactuated SLL versus the theoretical simulation derived kinematic curve.

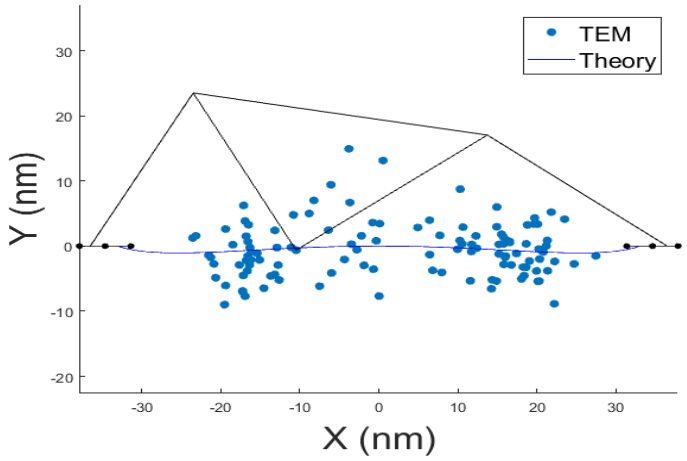


Figure 21: Theoretical kinematic model of SLL with experimental TEM data points measures the motion of p3, the triangle tip of SLL.

Measuring the motion of the theoretical and experimental SLL without actuation provides a model and control when analyzing the actuated SLL.

3.2: SLL actuation

The SLL is designed to be actuated via overhangs connected to l2, l5, l6, and l4 (lengths shown in Figure 19). The motion of p3 to the left or right side of the structure is actuated by combining unactuated SLL structures containing overhangs with an excess of complementary overhang staples. The unactuated SLL was folded with overhang sites consisting of 12 nucleotide long ssDNA, exposed to 10x excess of ssDNA actuating staples, and then placed in an incubator at 37°C for 2 hours.

The overhangs on the SLL for actuation have two separate configurations for binding – binding configuration 1 and 2. In figure 22, the joints corresponding to the location of the actuating overhangs and the binding configurations are shown.

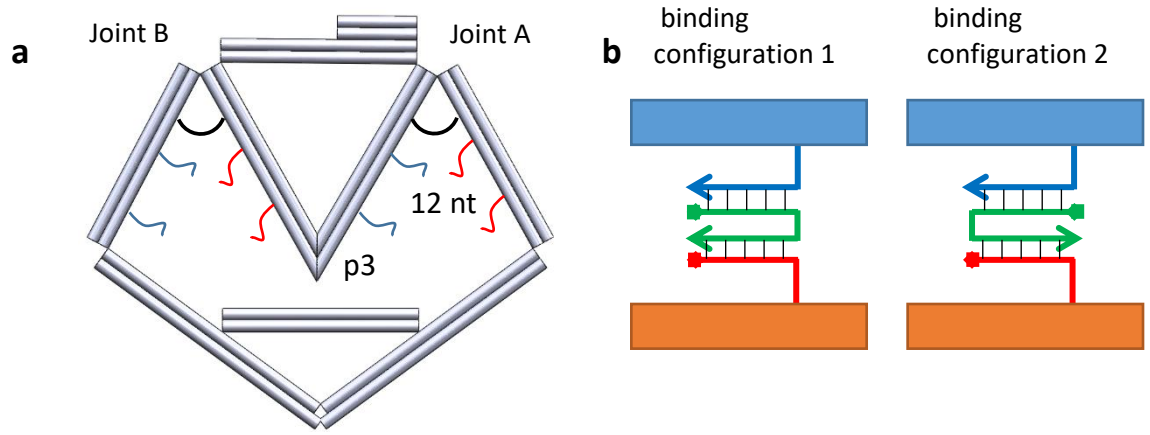


Figure 22: (a) SolidWorks schematic of SLL with overhangs (red and blue) that are 12 nucleotides (nt) long. The overhangs can bind to complementary ssDNA strands that close p3 to either Joint A or Joint B. (b) two different closing strand binding configurations, 1 and 2, for the complementary ssDNA that binds to the overhangs on the SLL. The different binding configurations allow for varying degrees of motion of p3.

The two different closing routings were tested for Joint A and Joint B, showing movement to either the left side or right side of the SLL. Since the SLL is a symmetric structure, there is a 22bp marker on the Joint A side to distinguish whether Joint A or Joint B was actuated when taking data.

Verification of actuated data was accomplished with the use of TEM, and a MATLAB code to calculate the probability of the movement of p3, shown in Figure 23.

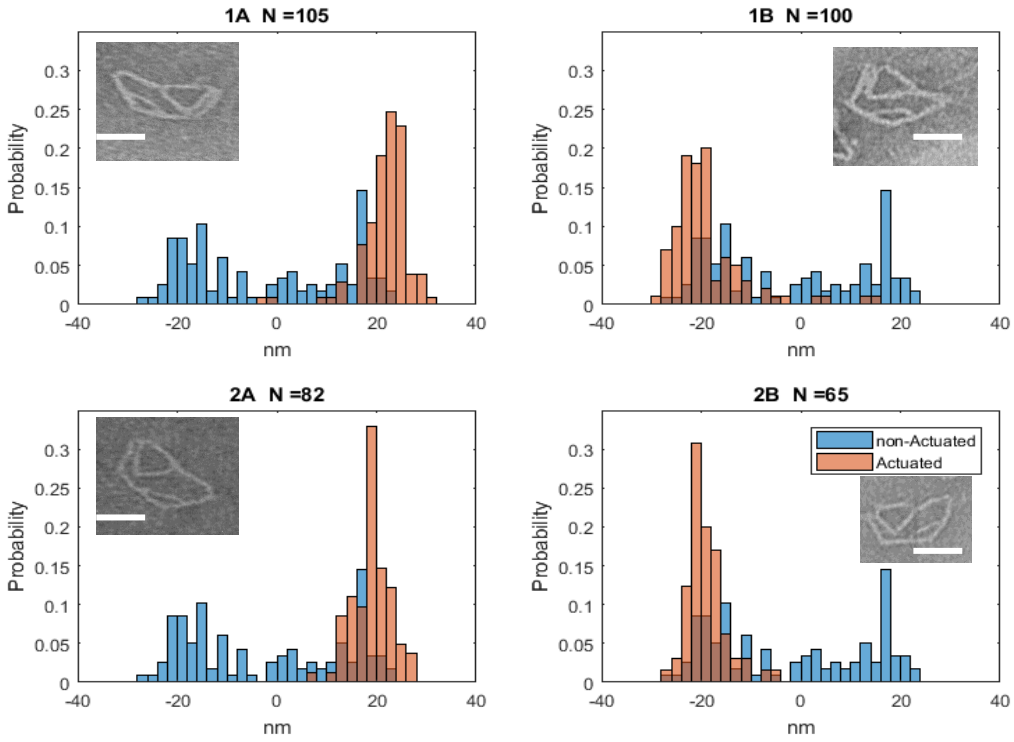


Figure 23: Raw data for actuation of SLL via closing binding configuration 1 and 2 for Joint A and Joint B. TEM images with scale bars 100nm are in the corners of each Joint actuation. The probability for actuation (closing) for each binding configuration is measured.

The probability data of the triangle tip p3 in the x-direction shows that binding configuration 2 for both Joint A and Joint B is higher than binding configuration 1. The different binding configurations show varying probabilities due to binding affinity.

3.3: SLL Conclusion

The SLL consists of multiple bundles that have measurable angles. With measured angles, the motion path of p3 can be calculated. The method of strand displacement is used to actuate p3 from the left and the right, allowing for measurements of the variability in motion of p3. As shown in figure 21, the theoretical motion path of p3 is plotted against experimental data from TEM. There is variability in the experimental and theoretical results. Near future work concerns characterizing the uncertainty of this motion.

Chapter 4: 4-Bar Linkage (4-Bar)

4.1: 4-Bar Linkage Design and Motivation

Combining the ideas of actuation and conformational changes, the 4-Bar Linkage is a structure that can transform from one conformation to another with the use of strand displacement and strand replacement. The concept of strand displacement was used by Nuemann et al, with a molecular machine made of DNA and has since been used in our lab as a way of actuating a structure (8,12). More recently, Song et al, has shown that triggered assemblies via triggering strands can change conformations, and relay information to a larger array of nanostructures (9). Relaying information, to an immune system for example, is a promising characteristic for an application such as a diagnostic or detection tool. Proof of concept for actuation (SLL) and proof of concept for triggering a hierarchical assembly (SLUG) is the basis of this work, thus combining these concepts allows for a dual mechanism. The transformable 4-Bar Linkage has 4 different configurations – A, B, C and D configurations.

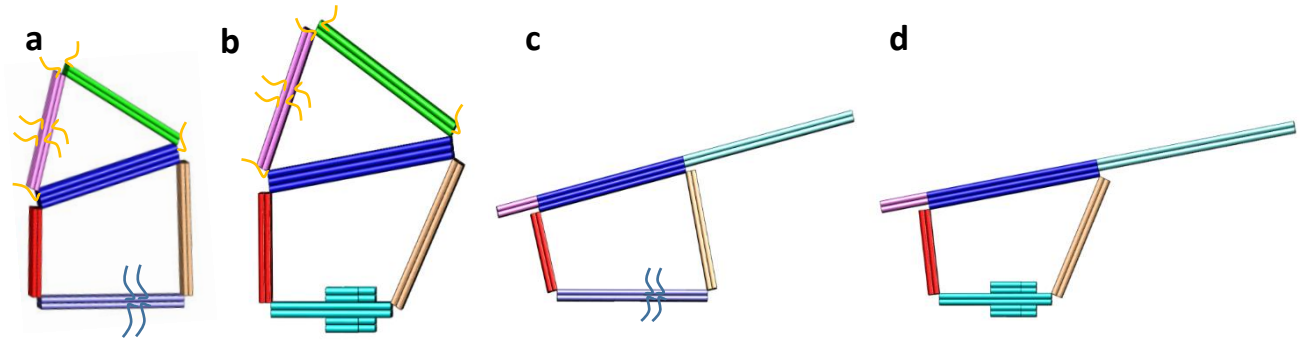


Figure 24: 4-Bar SolidWorks Schematics. (a) A configuration, with a triangle top and long base. (b) B configuration, with triangle top and short base. (c) C configuration with straight top and long base. A, B, and C configurations have toe-hold regions (ssDNA) for strand displacement and replacement. (d) D configuration, with straight top and short base. D is final configuration.

The core of each conformation remains the same, however the top and bottom portions of each configuration contain different staples, allowing for transformation between

configurations. Configuration A has a version that has overhangs used as a toe-hold, (ATH) allowing for breakage of the bundle via strand displacement.

Validation of well folded structures is shown with direct 2.5 day folding ramps are via gel electrophoresis and TEM images in figure 25. The band shift on the gel image shows that the various configurations run at a different length.

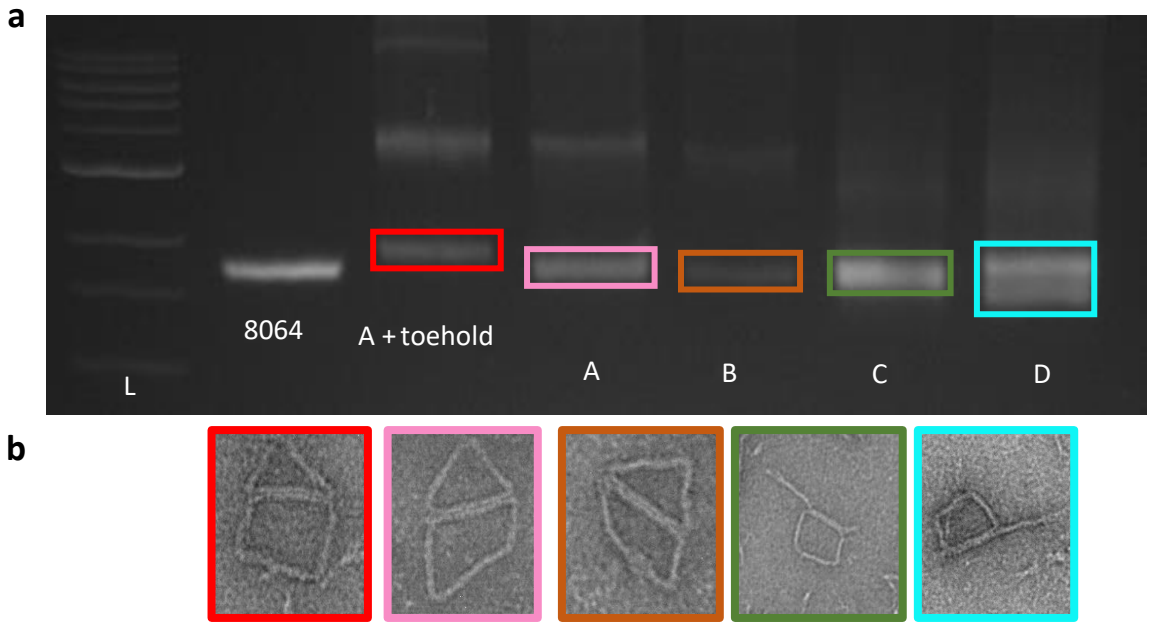


Figure 25: (a) left to right on gel image: 1 kb ladder, 8064 base scaffold, A+toehold, A configuration, B configuration, C configuration, and D configuration. (b) TEM images of each configuration, color coordinate to band on gel in (a).

Using directly folded ATH, three different pathways to transform from ATH to D configuration can be mapped out.

4.2: 4-bar Linkage Transformation Process

Transforming the 4-bar Linkage to different conformation consists of displacing certain staples, and replacing staples. Configuration A toe-hold (ATH) can be “broken” via strand displacement and have staples replaced to change conformations from ATH to B, ATH to C, and ATH directly to D configurations (Figure 24).

In figure 26, the pathways between a stable conformation and transition states are shown. The process schematic shows three different pathways which originate at ATH and terminate at a final structure of D configuration. The 4-Bar transformation process begins with ATH for each pathway. The addition of displacement and then replacement strands transforms ATH to a second configuration. The second configuration depends on which pathway is being constructed and can be seen in Figure 26. Two of the pathways have two transformation mechanisms containing displacement and replacement steps between each transformation mechanism. Between these transformation mechanisms, there is a transition state where toe-hold regions are displaced and replacement staples have not been added. Figure 26 summarizes the three pathways for transformation with the initial state ATH.

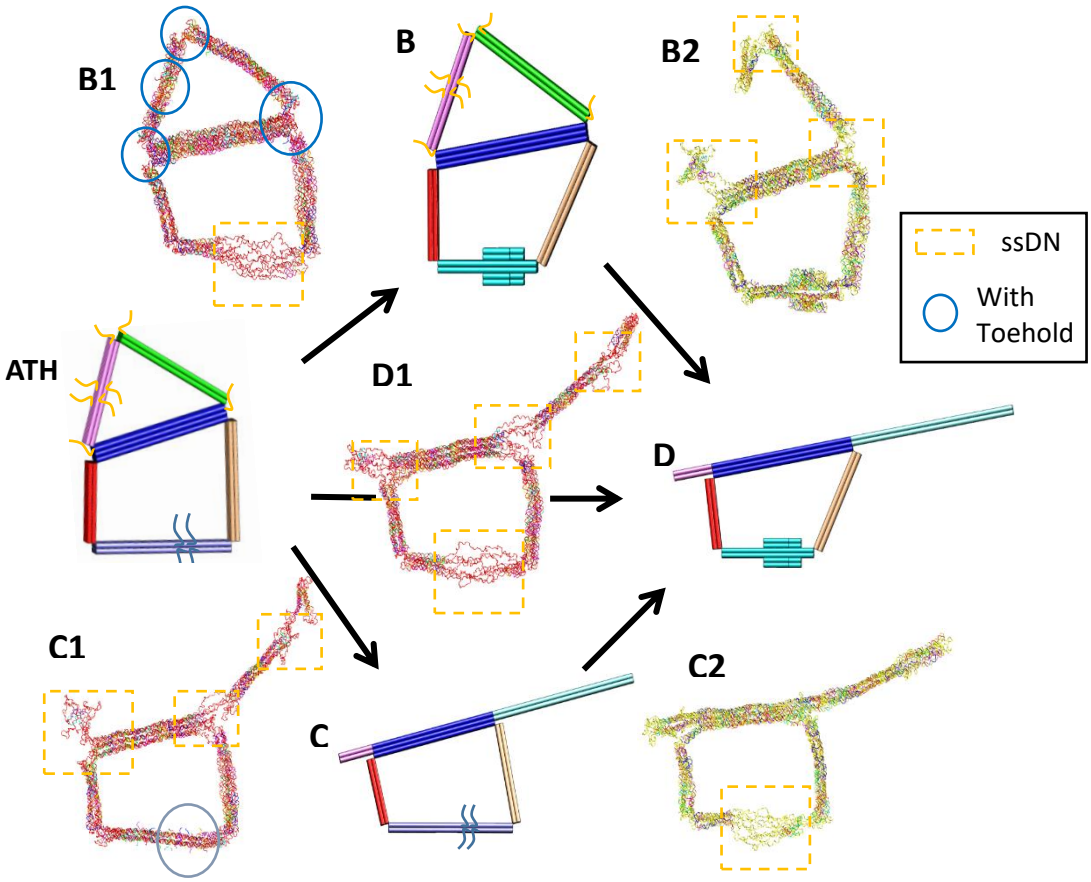


Figure 26: Transformation pathways. ATH is the starting point for each transformation pathway. B1 is a transition state, between ATH and B configurations, after strand displacement, with ssDNA and toehold regions for strand replacement. B2 is a transition state between B and D configurations, after strand displacement with ssDNA and toehold regions for strand replacement. C1 and C2 are like B1 and B2, however, for the C configuration transition states between ATH and C configurations and C and D configurations, respectively. D1 is the transition state between ATH and D configurations after strand displacement with ssDNA regions for strand replacement.

The transformation pathways are used as actuation mechanisms to change conformations. Changing conformations can be used to bind to different proteins or as a detection device. Since the 4-Bar has multiple configurations there are multiple applications for each configuration. For example, transforming from ATH with the triangle top to C mode with a straight top can be a potential mechanism for drug delivery. The triangle top can be used as a type of cryptic binding and reveal a payload once transformed to a straight top configuration such as C or D configuration.

Chapter 5: Methods

5.1: Thermal Annealing and Folding

Referencing "A primer to scaffolded DNA origami" (10), a structure is designed in caDNAno with a specific scaffold length appropriate to fold the structure. Design-specific staples are combined in five to ten-fold excess relative to the scaffold sequence, and combined with an ion-rich buffer ($MgCl_2$), as well as water and a buffer solution. The mixture is subjected to a thermal denaturation at 65°C, then lowered to a thermal annealing temperature, where staples anneal to the scaffold. Longer staples are typically placed on the inside of the scaffold routing. After thermal annealing, the mixture is then cooled to 4°C, creating a kinetic trap, preventing the structure to fold any further.

5.2: Purification Methods

Agarose gel electrophoresis is a method used to purify proteins, specifically by molecular weight. Gel purification is used in this lab to remove excess staples from a folding reaction, and to check for well-folded structures. A typical gel has multiple wells starting with one kb ladder, scaffold, and thermally annealed structure. A well-folded structure will run parallel or faster than the scaffold band, and a misfolded structure will run slower than the scaffold band. Secondary bands can be present if there are multiple configurations for a specific structure, however a misfolded structure will run slower than the scaffold, and the well-folded structure in the dimer band will be running parallel to the scaffold. For the work in this thesis, agarose gel electrophoresis was used to verify well-folded structures.

Another purification method used in this thesis work is polyethylene glycol (PEG) centrifugation. A solution containing equal volumes of PEG and structure are placed in an Eppendorf tube, spun in a centrifuge at 16000 G's for 25-30 minutes. The aqueous supernatant

is then removed, leaving a pellet of concentrated structure at the bottom of the Eppendorf tube. An aqueous buffer containing ions in the form of $MgCl_2$ is used as a resuspension buffer, diluting and breaking the structural pellet. PEG purification was used for the SLUG opening mechanism, as well as purification for SLL before actuation.

5.3: Transmission Electron Microscopy

Transmission Electron Microscopy (TEM) is a microscopic method used to collect images on the nanoscale. Structural samples are placed on a copper mesh grid with a carbon film, and negatively stained with 2% uranyl formate (UFo) as preparation for imaging. The samples are in a fixed conformation on the copper mesh grid, thus allowing for imaging on the microscope. A Tecnai G2 bioTWIN TEM was used courtesy of the Campus Microscopy & Imaging Facility (CMIF) at OSU, to image the SLUG, SLL, and 4-Bar Linkage for all experimental results.

Chapter 6: Conclusion and Future Work

6.1: Conclusion

DNA origami nanostructures have a variety of applications including biosensing and detection, drug delivery, nanomanufacturing. The work provided in this thesis aimed to show triggered assembly of different DNA origami structures, for future applications in detection, delivery, and capture of proteins. The structures were designed in an inspired manner from previous structures designed in the Nanoengineering and Biodesign Lab, such as the revolute joint (hinge), with several modifications that are tailored for future applications. The complex nature of these structures, with cryptic overhangs (SLUG), binding configurations for actuation (SLL), and toe-hold strand displacement sites (4-bar Linkage), allow for triggered assembly of hierarchical structures as well as conformational changes. Further work with these structures can ultimately lead to detection devices with specific trigger molecules, delivery vehicles with exposed binding sites, and capture tools to capture certain proteins, such as beta-amyloids, via polymerized structures.

6.2: Future Outlook for Triggered Assembly

The future applications of the SLUG design include creating a “net” of polymerized SLUG monomers, allowing for the capture of proteins or harmful molecules. Triggering an initially closed SLUG monomer to open and reveal internal overhangs or a therapeutic, can allow for the polymerization of the SLUG. The closing mechanism for the SLUG can be modified to use the latching strands instead of the internal binding sites.

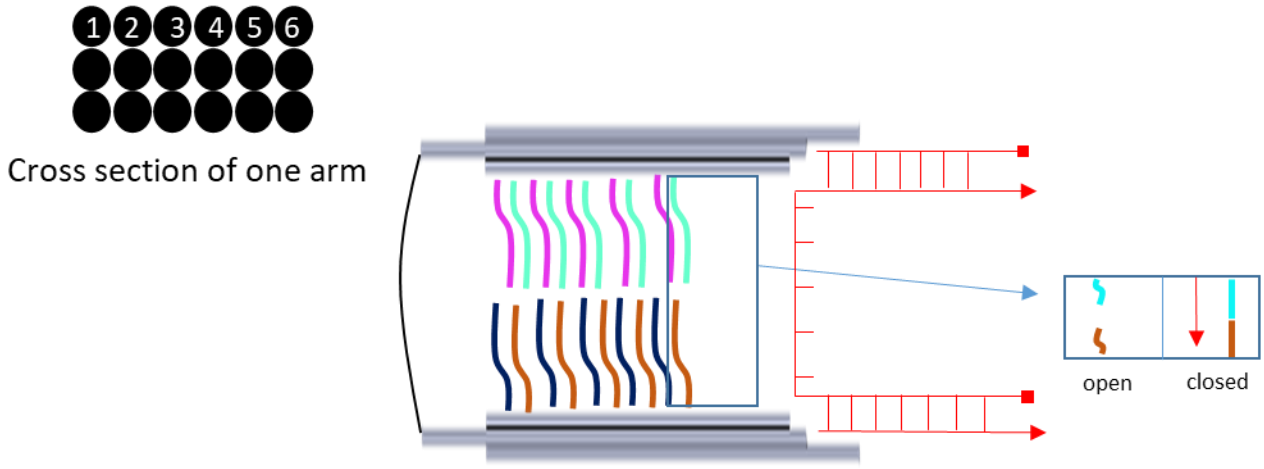


Figure 27: SolidWorks schematic of closed SLUG monomer with latching strand closing mechanism. The cross section of one arm of the SLUG shows 6 potential bundles for latching strand locations. Multiple latching strands can be added to close the SLUG ultimately varying the closing efficiency.

The number and location of the latching strands on the cross section of each arm can be changed depending on the desired accessibility of the internal overhang sites. For less accessible closed monomers, one latching strand on each bundle of the cross section of each arm (12 total latching strands) is necessary. Using the latching strands for a closing mechanisms allows the internal overhang sites of the SLUG to carry potential therapeutic payloads.

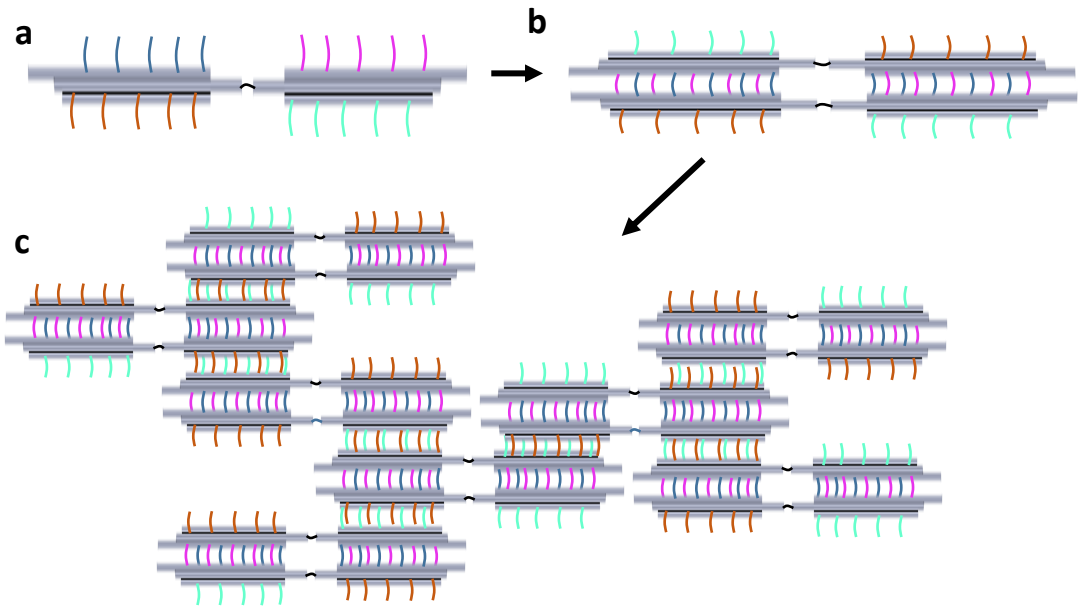


Figure 28: Hierarchical assembly of SLUG to create a net. (a) 1-Dimensional SLUG with external and internal overhangs for triggered assembly. (b) 2-Dimensional SLUG with external overhangs complementary to external overhangs on another SLUG monomer. (c) 3-Dimensional SLUG creating a net-like structure with tiling patters.

This trigger response can ultimately create a larger response in a physiological environment, such as an immune response, creating a diagnostic tool. Closed SLUG monomers can hide therapeutic payloads, and reveal these payloads upon a triggering mechanism. The applications of a triggered assembly can span farther than a diagnostic tool or a delivery system, the assembly can also be used to capture targeted proteins. For example, the exposed internal overhangs can attach a biotinylated molecule that can act as a therapeutic – such as vitamin C. Vitamin C is a known to accumulate in the immune system and facilitate immune cells to perform their task (14). There is research showing that the oxidized form of vitamin C can cross the blood brain barrier and be used as a therapeutic (13).

The actuation of the SLL can be applied for multiple future applications. One specific application that has a promising outlook is a signaling molecule. If a specific Joint (A or B) is

actuated, then a specific response can be applied. For example, using ion concentration gradients to trigger response mechanism is one possible outlook. There is current work being done in NBL that shows promising results for actuation with varying salt conditions and strand displacement. This can be applied by actuating a structure with hypertonic solutions then triggering the actuation of a revolute joint or for the SLL's multiple revolute joints.

The transformable properties of the 4-Bar Linkage have similar applications to the SLUG monomer. A future goal for the 4-Bar Linkage includes unlimited reversible transformation pathways. For example, the revolving Vernier mechanism has a controllable size of linear homomultimers (15). DNA origami nanostructures consisting of hollow cylinders and a rotatable shaft connected inside can stack with each other by the shape complementarity at the top and bottom surfaces of the cylinder (15). Control of the size of the homomultimer can be done by the limited twist of the angle of the rotatable shaft (15).

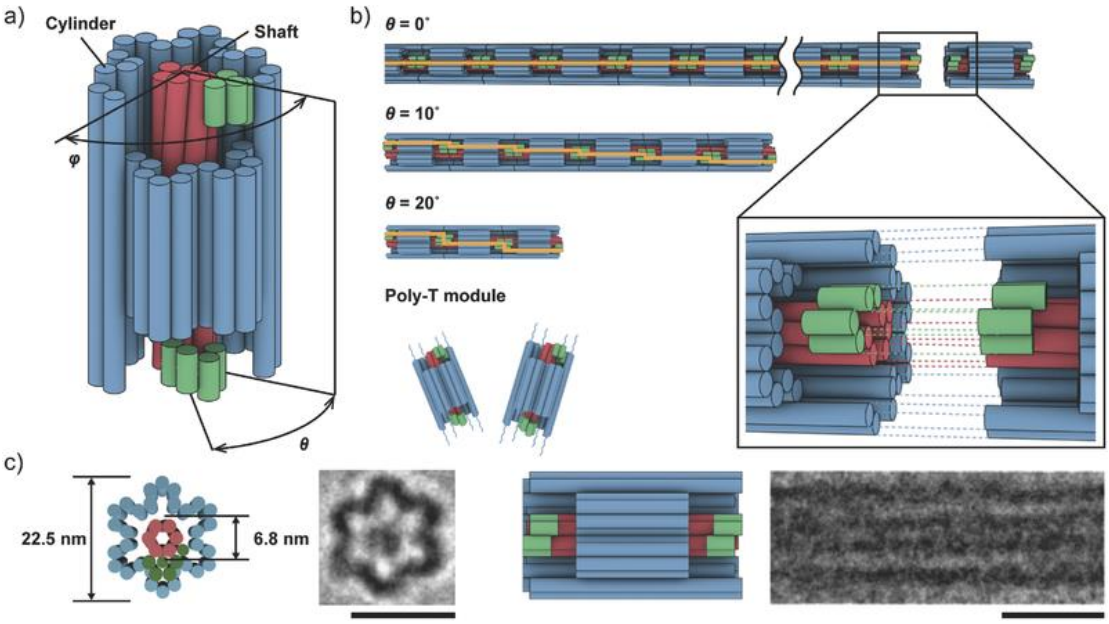


Figure 29: Revolving Vernier Mechanism Controls Size of Linear Homomultimer, Small 2017. The limits of the rotational motion and angle of the shaft place limits on the size of the homomultimer (15).

The 4-Bar can potentially follow a similar mechanism with each transformation causing a twist in the structure. This may allow for limitless amount of possible transformations or can place limitations or boundaries on the number of possible transformations.

After a transformation, each new configuration may have a different application or function. The triggered assembly can have multiple levels of sensitivity with varying functions and applications. For example, the final conformation or maximum level of sensitivity may function as a delivery system and deliver a payload. The transition state of conformation state before the final conformation may have a function such as a detection or diagnostic tool. The triangle top portion of the A configuration may be used as a cryptic binding site for a therapeutic. When triggered to configuration B, the triggered assembly may send a signal, and then at the final configuration the therapeutic may be released. Transforming mechanisms have a broad range of applications such as delivery systems and diagnostic tools.

Inspired by the 4-Bar Linkage and the SLUG, the 6-Bar has a triggered mechanism that can transform the initial structure to multiple conformations.

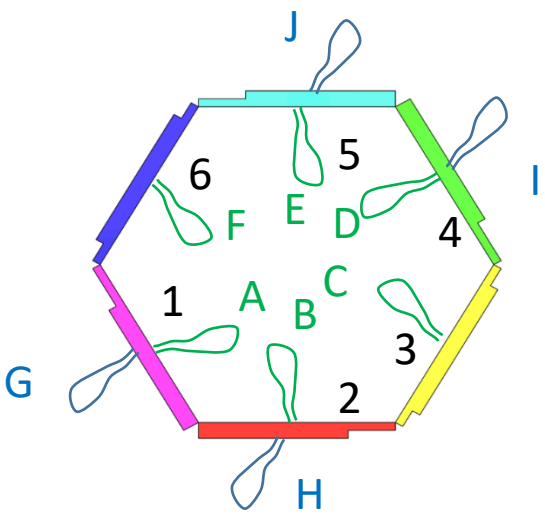


Figure 30: 6-Bar Linkage schematic. 6 bundles connected by ssDNA connections with overhang loops connected to each bundle.

Future work includes transformation of the 6-Bar Linkage to multiple conformations with different pathways like the 4-Bar Linkage. The 6-Bar Linkage is also capable of polymerizing because of the overhang loops. This type of polymerization is like the polymerization of the SLUG with the net, creating applications consistent to the SLUG.

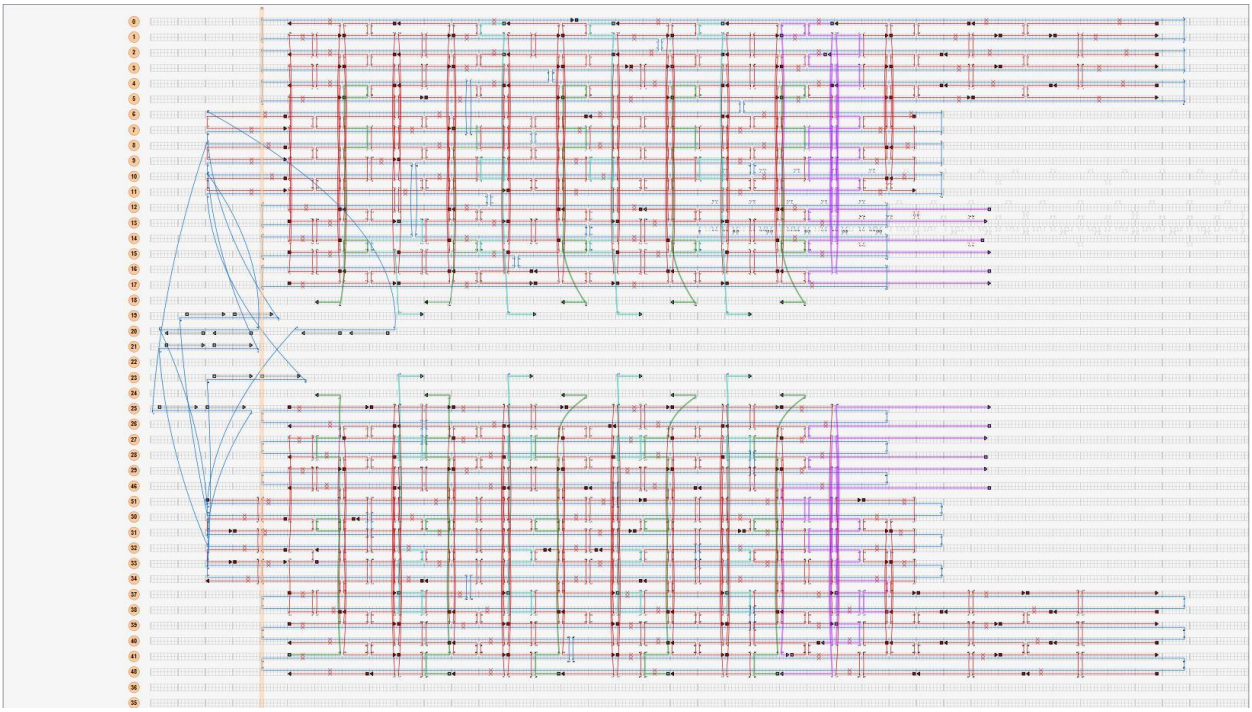
References

1. *Alzheimer's Disease & Dementia. Alz.org*
2. *Zhao, Zhao, et al. "*Organizing DNA Origami Tiles into Larger Structures Using Preformed Scaffold Frames." *Nano Letter*, vol 11, no. 7, 2011, pp. 2997-3002., doi: 10.1021/nl20163a.
3. Rothemund, Paul W.K. "Folding DNA to create nanoscale shapes and patterns." *Nature,* vol. 440, no. 7082, 2006, pp. 297-302., doi: 10.1038/nature04586
4. Douglas, Shawn M., et al. "Self-Assembly of DNA into nanoscale three-Dimensional shapes." *Nature,* vol. 459, no. 7245, 2009, pp. 414-418., doi: 10.1038/nature08016.
5. Sekiguchi, J., et al. "Resolution of Holliday junctions by eukaryotic DNA topoisomerase I." *Proceedings of the National Academy of Sciences,* vol. 93, no. 2, 1996, pp. 785-789., doi: 10.1073/pnas.93.2.785.
6. Fu, Tsu Ju, and Nadrian C. Seeman. "DNA double-Crossover molecules." *Biochemistry,* vol. 32, no. 13, 1993, pp. 3211-3220., doi: 10.1021/bi00064a003.
7. Marras, Alexander E., et al. "Programmable motion of DNA origami mechanisms." *Proceedings of the National Academy of Sciences,* vol. 112, no. 3, May 2015, pp. 713-718., doi: 10.1073/pnas.1408869112.
8. B. Yurke, A.J. Turberfield, A.P. Mills Jr., F.C. Simmel and J.L Nuemann, 'A DNA-fueled molecular machine made of DNA, " *Nature,* pp.605-608, 2000.
9. Song, Jie, et al. "Reconfiguration of DNA molecular arrays driven by information relay." *Science*, vol. 357, no. 6349, 2017, doi: 10.1126/science.aan3377.
10. Castro, Carlos Ernesto, et al. "A primer to scaffolded DNA origami." *Nature Methods,* vol. 8, no. 3, 2011, pp. 221-229., doi: 10.1038/nmeth.1570.
11. "Structural Biochemistry/Structural DNA Nanotechnology." *Structural Biochemistry/Structural DNA Nanotechnology – Wikibooks, open books for an open world.*
12. Su, Hai-Jun, et al. "The Kinematic Principle for Designing DNA Origami Mechanisms: Challenges and Opportunities." *Volume 5B: 39th Mechanisms and Robotics Conference,* Feb. 2015, doi: 10.1115/detc2015-46833.
13. Dsw. "Dehydroascorbic acid, a blood-Brain barrier transportable form of vitamin C, mediates potent cerebroprotection in experimental stroke." *Journal of Neurosurgical Anesthesiology,* vol. 14, no. 1, 2002, p.82., doi: 10.1097/00008506-200201000-00017.
14. 'Vitamin C and contribution to the normal function of the immune system: evaluation of a health claim pursuant to Article 14 or Regulation (EC) No 1942/2006. "*EFSA Journal*, vol. 13, no. 11, 2015, p.4298., doi: 10.2903/j.efsa.2015.4298.
15. Uchida, Takeo, et al. "Revolving Vernier Mechanism Controls Size of Linear Homomultimer." *Small,* vol. 13, no. 41, Dec. 2017, p.1702158., doi: 10.1002/smll.201702158.
16. Seeman, Nadrian C. "DNA in a material world." *Nature,* vol. 421, no. 6921, 2003, pp. 427-431., doi: 10.1038/nature

17. “DNA nanotechnology.” *Wikipedia*, Wikimedia Foundation, 12 Oct. 2017, en.wiipedia.org/wiki/DNA_nantechnology#Strand_displacement_cascades.
18. Zhou, Lifeng, et al. “Direct Design of an Energy Landscape with Bistable DNA Origami Mechanisms.” *Nano Letters,* vol 15, no. 3, 2015, pp. 1815-1821., doi: 10.1021/nl5045633.
19. Le, Jenny V. et al. “Probing Nucleosome Stability with a DNA Origami Nanocaliper.” *ACS Nano,* vol. 10, n.7. June 2016, pp. 7073-7084.,doi: 10.1021/acsnano.603218.
20. Stahl, Evi, et al. “Facile and Scalable Preparation of Pure and Dense DNA Origami Solutions.” *Angewandte Chemie*, vol. 126, no. 47, 2014, pp. 12949-12954., doi: 10.1002/ange.201405991.
21. Sharma, Rahul, et al. “Characterizing the Motion of Jointed DNA Nanostructures Using a Coarse-Grained Model.” *ACS Nano*, Mar. 2017, doi: 10.1021/acsnano.7b06470.

Appendix A: caDNAno files

A.1 caDNAno file for SLUG 6-site latching strand overhangs



Appendix B: Sequences

B.1: SLUG latching strand sequences

Order	Sequence 5' to 3'		Complementary 5' to 3'		Latching Strand:
1a	TGTGGACCCGACTAGGACCC	20	GGGTCCTAGTCGGGTCCACA	20	CTTCGTGTGTGTGTGTGTCCACCTAGGGGGTCCTAGTCGGGTCCACA
1b	GGACACACACACACACGAAG	20	CTTCGTGTGTGTGTGTGTCC	20	
2a	CGTGCACACGAAGTGGACCC	20	GGGTCCACTTCGTGTGCACG	20	GGGTCCACTTCGTGTGCACGTGGTCCCGGGTCCTAGGACCTAGGGGG
2b	CCCCCTAGGTCCTAGGACCC	20	GGGTCCTAGGACCTAGGGGG	20	
3a	CACACACGGACCCGACTAGG	20	CCTAGTCGGGTCCGTGTGTG	20	GGGTCCTAGTCGGGTCCTAGTCTAGGGCCTAGTCGGGTCCGTGTGTG
3b	CTAGGACCCGACTAGGACCC	20	GGGTCCTAGTCGGGTCCTAG	20	
4a	TGTGCACGGACCCGACTAGG	20	CCTAGTCGGGTCCGTGCACA	20	CCTAGTCGGGTCCGTGCACAGCTAGGAGTCTAGTCCTAGTCGGGTCC
4b	GGACCCGACTAGGACTAGAC	20	GTCTAGTCCTAGTCGGGTCC	20	
5a	CACGGATTTGACTAGGATTT	20	AAATCCTAGTCAAATCCGTG	20	CCTAGTCTAGTCGGGTCTAGCCTAGAGAAATCCTAGTCAAATCCGTG
5b	CTAGACCCGACTAGACTAGG	20	CCTAGTCTAGTCGGGTCTAG	20	
6a	CCCCCTAGACTAGACTAGAC	20	GTCTAGTCTAGTCTAGGGGG	20	GTCTAGTCTAGTCTAGGGGGGCCTAGTTGTGCACACACACACACACG
6b	CGTGTGTGTGTGTGTGCACA	20	TGTGCACACACACACACACG	20	

B.2: SLUG overhang strand sequences

AACCATCCGAAAGGAGCGTCCAACATAACGCTCATCAGTCATAAAG????????? **ATTGGTTG, ACTGTCCA,CAGTGATC**
AACCGTCTGTAGCGGTGGGGGTAAAGAGCAAATTCTCCGTATTTAAA???????? **ATTGGTTG, ACTGTCCA,CAGTGATC**
AGAACGTGTTAATGAACGAGAGGCGACGACGACATTAAATAAATTCGC???????? **ATTGGTTG, ACTGTCCA,CAGTGATC**
GTTCCAGGGTTTGCGCGGATAACGCCATGTTGTCTGGCCATTTTTT???????? **ATTGGTTG, ACTGTCCA,CAGTGATC**
TACCAGTCCCAGAGCTGAAAAGGTAATTAGCA???????? **ATTGGTTG, ACTGTCCA,CAGTGATC**
GAGGGAGGCAAAAGAAAGCGAACCATTGCTGAAAGTAATTCCTTGCTT???????? **GATCACTG, CAACCAAT,TGGACAGT**
TTCATTAACTACGAAGTCAGGATGAGGTCATACATGTTCAGTACAT???????? **GATCACTG, CAACCAAT,TGGACAGT**
CTTGAGCCTTTCCATTGCAATAATTACCAGATATCAACAATTTCATT???????? **GATCACTG, CAACCAAT,TGGACAGT**
ATCACCAGGAGCCACCATGATTATTAAGAAAAATAATATAGCATGTA???????? **GATCACTG, CAACCAAT,TGGACAGT**
CTCCATGTCCCCCAGCTTCAAATATTTAAATGTCTGAGAATTTTCCC???????? **GATCACTG, CAACCAAT,TGGACAGT**

Appendix C: MATLAB

C.1: TIF to PNG cropping (boxing) code

This code is used to crop TEM images for post processing. This code was not written by me and I do not take credit for the creation of this code.

```
% Matlab program for analysis of hinge-nucleosome
constructs

clc, clear all, close all


box_size = 300; % must be even number
% open file to write data
% text_file = '26bpNucConstruct_1-200.txt';
% if exist(text_file)~=0 %#ok<EXIST>
%     text_overwrite = questdlg('File name exists.
Overwrite?','','Yes','No','No');
%     if strcmp(text_overwrite,'Yes')
%         fid = fopen(text_file,'wt');
%     else
%         errordlg('Change output file name','file name.')
%         error('change text file name')
%     end
% else
%     fid = fopen(text_file,'wt');
% end

%
fprintf(fid,'Xv\tYv\tXba\tYba\tXn\tYn\tXta\tYta\tAngle\n');
figure(1)
```

```
set(gcf,'Position',[50 50 400 600]) % setting position of
full image figure

figure(2)
set(gcf,'Position',[800 100 300 300])

keep_going = 'Yes';
n=1;
while strcmp(keep_going,'Yes')
%     file_name = uigetfile('.tif');
    file_name = uigetfile('.tif');
    D = imread(file_name); % image to read
    figure(1)
    imshow(D), hold on
    new_image = 'No';
    while
and(strcmp(keep_going,'Yes'),strcmp(new_image,'No'))
        figure(1)
        title('Select particle to box')
        clear xb_c yb_c
        [xb_c, yb_c] = ginput(1); % select point that will
be center of boxed particle
        plot(xb_c, yb_c,'ro') % plot picked point
        % points to draw box
        xb_min = floor(xb_c - box_size/2);
        xb_max = xb_min + box_size;
        yb_min = floor(yb_c - box_size/2);
        yb_max = yb_min + box_size;
        xb = [xb_min xb_max xb_min xb_min];
        yb = [yb_min yb_max yb_max yb_min];
        plot(xb,yb,'r') % plotting box around particle
        % getting new image that is just the boxed particle
        D_box = D(yb_min:yb_max,xb_min:xb_max,:);
        figure(2)
        imshow(D_box), hold on
        truesize(figure(2),[400 400])
        take_data = questdlg('Take
data?','','No','Yes','Yes');
        if strcmp(take_data,'Yes')
%             h1=impoly; % draw a polygon with 4 points
(vertex, end of bottom arm, nuc, end of top arm, vertex)
% %               pause
%             pos1 = getPosition(h1);
%             % Calculating hinge angle
```

```
%              v1 = [(pos1(2,1)-pos1(1,1)) (pos1(2,2)-
pos1(1,2))];
%              v2 = [(pos1(4,1)-pos1(1,1)) (pos1(4,2)-
pos1(1,2))];
%
cos_theta=dot(v1,v2)/(sqrt(v1(1)^2+v1(2)^2)*sqrt(v2(1)^2+v2
(2)^2));
%              theta = acos(cos_theta)*180/pi;
%              % writing data to text file
%
fprintf(fid,'%4.2f\t%4.2f\t%4.2f\t%4.2f\t%4.2f\t%4.2f\t%4.2
f\t%4.2f\t%4.2f\n',pos1(1,1),pos1(1,2),pos1(2,1),pos1(2,2),
pos1(3,1),pos1(3,2),pos1(4,1),pos1(4,2), theta);
             % logic to keep going
            keep_going = questdlg('Keep
going?','','No','Yes','Yes');
            new_image = questdlg('Select new
image?','','Yes','No','No');
%             text1 = text_file(1:(length(text_file)-4));
             text2 = file_name(1:(length(file_name)-4));
             box_image_file =
strcat(text2,'_',sprintf('%04s',num2str(n)),'.png');
             imwrite(D_box,box_image_file,'tif')
            n=n+1; % increment for saving images
        end
    end

end

%fclose(fid);

% impoly is polygon function
% pos = getPosition(h)
```

C.2: SLUG latching and overhang strand sequence generator

This code was not written by me and I do not take credit for the creation of this code.

```
clear all
clc
tic
%%
% specify the parameters of the overhang you wish to generate. Overhang
% length must be great than 5 and will be generated in multiples of 4. You
% may need to truncate your results
overhangLength = 20;
minGCcontent = 50;
```

```
maxGCcontent = 80;
solutions = 12; %will generate approximately this many solutions

% Import the scaffold sequence here
fileID = fopen('p8064.txt');
C = textscan(fileID,'%s');
fclose(fileID);
scaffoldtext = char(C{1});
% Import your staple list here (as a .txt file)
stapletext = importdata('Latching_strand_staple_list.txt','%s');
stapletext{numel(stapletext)+1}=scaffoldtext;
%% STEP 1: find all starting sequences
bases = categorical({'A' 'T' 'G' 'C'});

len = 4;
all = bases;
for i=1:len-1
    temp = all.*bases';
    all = reshape(temp,[1 numel(temp)]);
end

%% Convert to character strings
seqs = cell([1 numel(all)]);
for i=1:numel(all)
     seqs{i}=strrep(char(all(i)),' ','');
end

%% Evaluate all sequences for the number of times they are found in scaffold
goodseqs = zeros([1, numel(all)]);
scores = cell(1,numel(all));
for s=1:numel(stapletext)
    for i=1:numel(all)
        temp = strfind(stapletext{s}, seqs{i});
        goodseqs(i) = goodseqs(i)+numel(temp);
    end
end

%% Initialize with sequences that are not found in the scaffold
% Note you can adjust the number in quantile() to give you more or less
% seeded sequences
k =find((goodseqs<quantile(goodseqs,.05)));
round1 = categorical(seqs(k))
figure(1);
plot(goodseqs)

%% Append all possible bases to all starting sequences and only keep
sequences that appear the least in the scaffold
all = round1
for i=1:floor(overhangLength/4)
    added = 4
    for k=1:added
        temp = all.*bases';
        all = reshape(temp,[1 numel(temp)]);
    end
    seqs = cell([1 numel(all)]);
    for j=1:numel(all)
        seqs{j}=strrep(char(all(j)),' ','');
```

```
    end
    goodseqs = zeros([1, numel(all)]);
    for s=1:numel(stapletext)
        for j=1:numel(all)
            for k=1:length(char(seqs{1}))-4
                charseq = char(seqs{j});
                temp = strfind(stapletext{s}, charseq(k:k+4));
                goodseqs(j) = goodseqs(j)+numel(temp);
            end
        end
    end
    k =find(goodseqs<=quantile(goodseqs,solutions/numel(all)));
    finalists = categorical(seqs(k))
    gccont = zeros([1 length(finalists(:,1))])
    for i=1:length(finalists)
        temp = oligoprop(char(finalists(i)));
        gccont(i) = temp.GC;
    end
    g = find((gccont<maxGCcontent)&(gccont>minGCcontent));
    all=categorical(finalists(g))
    figure(1);
    plot(goodseqs)
    pause(0.1);
end

%% Show alignments
fprintf('Total of %d solutions for overhangs found\n',numel(all));
fprintf('%s\n',string(all(:)));

%%
fprintf('\n\n the following show where each overhang has the greatest
affinity based on swalign');
pause();


for i=1:numel(all)

    [score align]=swalign(char(stapletext(end)),char(all(i)))
    pause();
end
```

C.3: 5-point analysis code for SLL

This code was not written by me and I do not take credit for creating this code.

```
% Matlab program for analysis of hinge-nucleosome
constructs
% This pulls from individual tiles of boxed structures,
specifically
% hinges, allows you to pick coordinates (4pts) of vertex,
end of bottom
% arm, nucleosome, end of top arm, and vertex again.  Will
keep the
```

```
% coordinates and angle measurement as comment within a new
renamed file.

% make sure directory is proper at line 117 ish
% make a "Down Folder"
% This has a tendency to error out due to dimension issues.
Please save in
% an excel file
% It doesn't actually end and gets funky everytime it hits
a multiple of 10

clc, clear all, close all
%clc, close all
  folderName='NewPng';
if isdir(folderName)
% rmdir(strcat(pwd,'\',folderName) );
rmdir NewPng s
end

%work with V3


k=1;
n=1;
m=1;

% open file to write data
text_file = 'DataWithImgName.txt';
if exist(text_file)~=0 % #ok<EXIST>
    text_overwrite = questdlg('File name exists.
Overwrite?','','Yes','No','No');
    if strcmp(text_overwrite,'Yes')
        fid = fopen(text_file,'wt');
    else
        errordlg('Change output file name','file name.')
        error('change text file name')
    end
else
    fid = fopen(text_file,'wt');
end

fprintf(fid,'X1\tY1\tX2\tY2\tX3\tY3\tX4\tY4\tX5\tY5\tImage\
n');
```

```
% %loading images
% fname1=uigetfile('*.png','Select first data files for
averaging. ');
% fname2=uigetfile('*.png','Select last data files for
averaging. ');
%
% doubledouble = questdlg('Is the last file in double
digits?','','No','Yes','Yes');
%         if strcmp(doubledouble,'Yes')
%             seq1=fname1((length(fname1)-
7):(length(fname1)-6)); %first file (raw file)
%             seq2=fname2((length(fname2)-
8):(length(fname2)-7)); %last file
%         else
%             seq1=fname1((length(fname1)-
7):(length(fname1)-6)); %first file (raw file)
%             seq2=fname2((length(fname2)-
7):(length(fname2)-6)); %last file
%         end
%
% s1n=str2double(seq1);
% s2n=str2double(seq2);
%
%
% % d = uigetdir(pwd, 'Select a folder');
% % subimages = dir(fullfile(d, '*.jpg'));
%
%
%
% % Ns=s2n-s1n+1;
% % Ns=300; %Total number of tiles
%
%   for aa=s1n:s2n
%     if aa<10
%       i_strs=sprintf('%1s',num2str(aa)); %Substitute %02s
for 10+ images
%         for j=0:9 %Number of tiles per file, adjust as
necessary.  Will have lag between photos.
%     i_str=sprintf('%02s',num2str(j));
%     fname{m,:}=strcat(fname1(1:(length(fname1)-
8)),i_strs, '.',i_str,'.png');
%     m=m+1;
%         end
%     else
```

```
%     i_strs=sprintf('%02s',num2str(aa)); %Substitute %02s
for 10+ images
%         for j=0:9 %Number of tiles per file, adjust as
necessary.  Will have lag between photos.
%     i_str=sprintf('%02s',num2str(j));
%     fname{m,:} =strcat(fname1(1:(length(fname1)-
8)),i_strs, '.',i_str,'.png');
%     m=m+1;
%         end
%     end
%
%     aa=aa+1;
%   end
%---------------------------------CM

d = uigetdir(pwd, 'Select a folder');
subimages = dir(fullfile(d, '*.png'));

fname=cell(size(subimages,1),1) ;
for kw=1:length(fname)
  fname{kw}=subimages(kw).name ;
end

Pstar_rules = char(fname);
Ns= length(fname);
figure(2)
set(gcf,'Position',[800 100 300 300])

while k<=Ns
    fname3(1,:) = Pstar_rules(k,:)
    fname4 = exist (fname3);

    if fname4 == 0
        %File does not exist, skip to the bottom of loop
and continue
          k = k+1;
          print('a.txt')
    else
        C = imread(Pstar_rules(k,:)); % image to read
%       D = padarray(C,[50,50], 'replicate','both');
        f2=figure(2)
%       imshow(D_box), hold on %for galleries
        imshow(C), hold on %for single tiles
        truesize(figure(2),[400 400])
```

```
        take_data = questdlg('Take
data?','','No','Yes','Yes');
        if strcmp(take_data,'Yes')
            h1=impoly; % draw a polygon with 4 points
(vertex, end of bottom arm, nuc, end of top arm, vertex).
            %Start with vertex and towards the end of the
bottom arm
            %pause
            pos1 = getPosition(h1);
            % Calculating hinge angle
            v1 = [(pos1(2,1)-pos1(1,1)) (pos1(2,2)-
pos1(1,2))];
            v2 = [(pos1(4,1)-pos1(1,1)) (pos1(4,2)-
pos1(1,2))];
%
cos_theta=dot(v1,v2)/(sqrt(v1(1)^2+v1(2)^2)*sqrt(v2(1)^2+v2
(2)^2));
%             theta = acos(cos_theta)*180/pi;
            text(pos1(1,1),pos1(1,2),'1st' )

            dfgdg=45;


            Frame = getframe(f2) ;
%             figure ; imshow(Frame.cdata);


            if ~isdir(folderName)
            mkdir(folderName);
            end
            filename=strcat(pwd,
'/',folderName,'/','New',Pstar_rules(k,:)) ;
            imwrite(Frame.cdata,filename)

            % Writing data to text file

fprintf(fid,'%4.2f\t%4.2f\t%4.2f\t%4.2f\t%4.2f\t%4.2f\t%4.2
f\t%4.2f\t%4.2f\t%4.2f\t%s\n',pos1(1,1),pos1(1,2),pos1(2,1)
,pos1(2,2),pos1(3,1),pos1(3,2),pos1(4,1),pos1(4,2),pos1(5,1
),pos1(5,2),Pstar_rules(k,:)  );

            % Write image file
%             text2 = Pstar_rules(k,:)
%                 if k == 1
```

```
%                     save_name= uiputfile('.png');
%                 else
%                 end
%             j_str=sprintf('_%003s',num2str(k));
%
box_image_file(k,:)=strcat(d,save_name(1:(length(save_name)
-4)),j_str,'.png');
%             %Change box_image_file directory as needed.
%             exnum = [pos1(1,1) pos1(1,2) pos1(2,1)
pos1(2,2) pos1(3,1) pos1(3,2) pos1(4,1) pos1(4,2) theta];
%             ex =num2str(exnum, 4) ;
%
%imwrite(D_box,box_image_file,'png','Comment',ex) %for
galleries
%             imwrite(C, box_image_file(k,:),'Comment',ex);
%for single tiles

            n=n+1; % increment for saving images

        end
        k = k+1;
    end

end


fclose(fid);
```